How scientists and physicians use Twitter during a medical congress

Muge Cevik, MD, MSc, MRCP(UK)[1,2,3], David S.Y. Ong, MD, PharmD, PhD [4,5] and Graham Mackenzie, MD [6]

Affiliations

[1]Division of Infection and Global Health Research, School of Medicine, University of St Andrews, St Andrews, United Kingdom

[2]Regional Infectious Diseases Unit, Western General Hospital, Edinburgh, United Kingdom

[3]Department of Laboratory Medicine, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom

[4]Department of Medical Microbiology and Infection Control, Franciscus Gasthuis & Vlietland, Rotterdam, the Netherlands

[5]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

[6]Penicuik Health Centre, Edinburgh, United Kingdom

Author of correspondence:

Name: Dr Muge Cevik

Address: Division of Infection and Global Health Research, School of Medicine, University of St Andrews, Fife, KY16 9TF

Telephone number: 07531993677
Email address: mc349@st-andrews.ac.uk

Key words: social media, Twitter, MedTwitter, science communication, medical conference

Abstract

Objectives: During medical congresses Twitter allows discussions to disseminate beyond the congress hall and reach a wider audience. Insights into the dynamics of social media interactions during congresses, dissemination of scientific information and the determinants of a successful tweet may allow us to better understand social media's role in science communication.

Methods: We retrospectively extracted social media data during the European Congress of Clinical Microbiology and Infectious Diseases (ECCMID) 2017 and 2018 using NodeXL. We compared social media activity during these two congresses. Subsequently, we conducted in-depth analyses to identify the components of a successful tweet and multivariable analysis to assess independent factors associated with retweet activity.

Results: In 2018, approximately 13,000 delegates attended ECCMID, but only 591 Twitter accounts actively tweeted about the congress. Although fewer tweets were posted in 2018 compared to 2017 (4,213 vs 4,657, respectively), ECCMID2018 generated a 63% increase in the total number of retweets ($p < 0.001$). According to multivariable logistic regression analysis, using multimedia, URL or hashtags and mentioning other Twitter account(s) were independently associated with retweet success. Mentioning of other users and use of multimedia were the only consistent predictors of retweets irrespective of the number of followers.

Conclusions: A substantial increase in retweet activity and a modest increase in the number of influential Twitter accounts were observed between two successive congresses. Dissemination

of scientific messages is more successful when connected accounts are actively involved in social media activity, and social media posts constitute the right combination of components.

Introduction:

Twitter is a microblogging platform used as a communication tool by scientists as a means to exchange scientific information and ideas, network with peers and initiate direct engagement with non-scientific audiences [1-3]. Use of Twitter during congresses has also gained momentum in helping important scientific discussions to go out of the congress hall and reach a wider audience [4]. Unlike conventional media, social media facilitates two-way interaction combining the roles of broadcaster and audience [5]. It also provides timely dissemination of knowledge and expert opinion with the virtual participants. Tweets can be shared (retweeted) and may include media (images and/or video), keywords which can be indexed and searched by the social media network (hashtags), mentions of other users, and links to other information on the web. These components can have an impact on the amplification of a tweet, which varies depending on multiple factors [6, 7].

The European Congress of Clinical Microbiology and Infectious Diseases (ECCMID) is the largest international congress in the field, organised by the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) which has over 7,000 individual and 30,000 affiliated members. ECCMID is a 5-day congress attracting nearly 13,000 scientists, physicians and other healthcare staff from all around the world [8]. Many of the topics discussed at this congress (e.g. antimicrobial resistance, emerging infections and vaccines) are of potential interest to other professionals who cannot attend the congress in person and the general public.

This study was performed: (1) to provide a descriptive overview of the social media activity during the 27th ECCMID 2017 and 28th ECCMID 2018 congresses; (2) to provide comparison between two successive congresses; (3) to identify the components of tweets associated with an increased probability of dissemination (retweeting).

Methods:

Study design and data extraction

We retrospectively extracted social media data (tweets, retweets, mentions) covering the congress days; 22-25 April 2017 for ECCMID2017 and 21-24 April 2018 for ECCMID2018 utilising NodeXL, which is an Excel add-on developed by SMRF [9]. The NodeXL extract was analysed to establish the number of Twitter accounts using the congress hashtag for the given year (#ECCMID2017 or #ECCMID2018). We identified original tweets using Tweet ID (a unique 19-digit identifier generated for each tweet) and retweets using the unique 19-digit Retweet ID [10], extracting username (Twitter handle) and mentions of other Twitter accounts from these tweets.

The NodeXL extracts were repeated and refined until data were available for the entire congress period, thereafter, duplicate tweets and commercial tweeters were excluded. Subsequently, for completeness, the number of tweets extracted by NodeXL was compared with the number of tweets identified via Twitter search on Google Chrome browser. NodeXL successfully captured 97.5% of tweets from the congress period. In total 282 "quoted tweets" and 259 replies from ECCMID2018 data were excluded prior to univariable and multivariable analysis, leaving 3,653 tweets. Components of tweets were identified by NodeXL. Hashtags were identified using the preceding "#" symbol, and Twitter handles using the leading "@" symbol. For the multivariable analysis tweets mentioning other tweets were identified using

Twitter.com domain and replies were identified as tweets beginning with the "@" symbol. Data extracted from Twitter is publicly available and is therefore exempt from IRB panel review. We have obtained opt-out consent from the Twitter users included in the sub-group analysis.

Data analysis

First, social media activity of ECCMID 2017 and ECCMID 2018 was compared using $\chi 2$ test and descriptive statistics. NodeXL extracts were mapped according to the estimates of connectedness using information from the Twitter Application Programming Interface (API), and outputs were combined into a single network map. NodeXL uses estimates of connectedness, ranking users by "betweenness centrality" which measures the number of "shortest paths" that pass through each Twitter user using network theory [11]. Subsequently, we classified users to identify key influencers using Venn diagrams.

We used the number of retweets as an outcome measure of the reach and influence of a tweet. We excluded replies to tweets and tweets that quoted another tweet and replies, because replies are only seen by the user replied to and mutual followers, unless searched for specifically, and the quoted tweet is typically displayed rather than any images in the tweet. In univariable analysis, we assessed possible predictors of retweets. Based on previous studies [6,7], variables of interest were: inclusion of multimedia (images and/or video), a link to other information on the web (Uniform Resource Locator (URL)), mention of other tweeter(s), the number of followers of a Twitter user and use of hashtags other than the congress hashtag. We then performed multivariable logistic regression analysis using ECCMID2018 data to identify independent predictors of retweet activity. We restricted this particular analysis to a single year to reduce bias as Twitter introduced new rules over time; the way the tweets are displayed, and the length of tweets changed in 2018. We tested for multicollinearity and interactions among

the included variables. Finally, we performed a subgroup analysis among the top Tweeters based on the number of retweets received. We performed a descriptive analysis of the number of followers, number of tweets and retweets received among these influencer users.

To ensure accurateness of the extract, the top 10 tweeters identified by betweenness centrality in the NodeXL report were compared with the top 5 Twitter accounts identified in each of the seven sections of the Venn diagram for tweeter, retweeter and mentioned users. Subsequently, the most popular tweets were selected for further analysis and collected into congress summaries listing individual tweets chronologically [11, 12]. The distribution of the tweeters included in the Congress summaries was also matched the Venn diagram output.

Results:

Table 1 displays the number of Twitter accounts involved, and tweets and retweets generated both in 2017 and 2018 Congress social media activity. Although the number of accounts engaged in active tweeting remained relatively static (590 vs 591) and fewer tweets were posted (4,213 vs 4,657) in 2018 compared to 2017, ECCMID2018 generated a 63% increase in retweet activity (total number of retweets). Besides, the proportion of accounts engaged in retweeting has increased ($p< 0.001$), with "just retweeters" making up the single largest group for both years. The median number of followers for active tweeters also increased from 203 (IQR 47-891) in 2017 to 278 (IQR 91-1,030) in 2018 ($p< 0.001$). The number of tweeters who received 80% of retweets was similar in both years (67 in 2017 vs 72 in 2018), while 222 (38%) active tweeters received zero retweets in 2017 vs 186 (31%) in 2018 ($p< 0.001$).

Mapping Twitter activity

Figure 1 illustrates the extent of connections and dissemination of tweets [14]. Each picture represents a Twitter user who either tweeted, were mentioned, and/or retweeted posts using #ECCMID2018. The most influential tweeters are situated at the centre of different groups, and the connecting lines are weighted by the strength of the connection [11]. The map displays a predominantly "tight crowd" pattern formed by highly connected tweeters (G1-G4). This very large and complex map has been included to demonstrate how connected the tweeters were during ECCMID2018.

Venn diagram (Figure 2) has allowed us to classify users into tweeters, retweeters and those mentioned and helped to distinguish the key users. Tweeter accounts in the central zone of the Venn diagram, representing 4,8% of all accounts, were actively tweeting and also were mentioned and retweeted. These were the most influential tweeters and as a result made up the largest group of tweeters in the Congress summary [13] and were included in the sub-group analysis for further evaluation. The number of influential accounts showed a modest 13% increase in 2018 (from 127 to 144). The "just retweeted" category was the largest group (72% of the active accounts in 2018), and the number of accounts in this category has substantially increased from 1,167 in 2017 to 1,904 in 2018 (63% increase). These tweeters did not generate their own content, but they were important in disseminating the information. According to the Congress summaries [12, 13], the "just mentioned" category included speakers at the Congress, international organisations like ESCMID, university departments, journals and journal editors and made up the second largest group of Twitter accounts on the NodeXL map.

Independent predictors for retweets

The components of a successful tweet were initially studied using univariable analysis based on ECCMID2018 data. Overall, out of 3653 tweets 86% (n=3158) included one or more additional components: 62% (n=2248) included an image (n=2163) or a video (n=85) which is categorised as multimedia, 38% (n=1392) mentioned other tweeter users, 35% (n=1281) used a non-conference hashtag, and 17% (n=634) included a URL. Fifty three percent of tweets (n=1930) were posted by tweeters with less than 1,000 followers. As shown in Figure 3, there was a stepwise increase in the number of retweets received with the inclusion of more components. Based on this analysis, use of multimedia, number of followers, URL, and mentions of other Twitter users were all positively associated with retweets and were selected for multivariable analysis.

The impact of individual components of tweets, which were identified by previous component analysis, was further studied by multivariable logistic regression analysis. The use of multimedia (images and/or video), mentioning of other tweeters, the use of other hashtags, the inclusion of an URL and the number of followers were all independently associated with retweets (Table 2). However, the number of followers had significant interactions with mention of other tweeters and also with inclusion of an URL. Despite inclusion of these interaction terms into the model, all individual tweet characteristics remained significant. Based on this observation a subgroup analysis was performed in tweeters with >1,000 followers and <1,000 followers. Mentioning other Twitter users and inclusion of multimedia were the only independent predictors of retweets in tweeters with fewer than 1,000 followers, whereas among tweeters with more than 1,000 followers all four tweet characteristics remained significant predictors of retweets.

Sub-group analysis

The top tweeters of the ECCMID2018 are shown in Supplementary table, which includes the top influencer ESCMID, as well as a mixture of personal accounts, organisation and journal accounts, and commercial companies. In this analysis, we compared the number of followers of the top tweeters, the number of tweets they have posted, and the number of mentions and retweets these accounts have received. ESCMID ranked the highest in each category. However, we observed a discrepancy among other tweeter accounts. For instance, some accounts with higher numbers of followers and tweets received fewer retweets than other accounts with a smaller number of followers and tweets.

Discussion

In this study, we observed a substantial rise in retweet activity and a modest increase in the number of influential Twitter accounts between two successive congresses. Although it is difficult to ascertain the precise motives behind this enhanced activity through this analysis, the findings indicate the importance of social media connections, which is displayed as "tight crowd" pattern on the NodeXL map formed by highly connected tweeters i.e. ESCMID account. Besides, during ECCMID 2018, the Trainee Association of ESCMID (TAE) steering committee members, all physicians, were involved in the congress social media planning and have actively tweeted from ESCMID and TAE twitter accounts, whereas during ECCMID 2017 mostly non-physicians were involved, which might have influenced the quality, content and appropriateness of the tweets. This emphasises the importance of involving physicians and scientists in outlining and implementing social media activity for medical congresses. Additionally, these results highlight that there is an evolving interest in social media during conferences. Many scientists indicate that they actively follow conferences remotely through Twitter [3]; however, as observed in this study the main generators of content remain limited. In 2018, approximately 13,000 delegates attended ECCMID, but only 591 Twitter accounts

actively tweeted about the Congress. In an era of widespread social media usage, there is a reluctance by some scientists and physicians to use this medium.

Social media posts generated during these congresses reached a broader audience. In comparison with a cardiology conference [15], ECCMID2018 had a more connected pattern of tweeting. Besides, the median number of followers for Twitter accounts increased between years. In a recent study on social media dynamics, scientists and physicians with over 1,000 followers have been shown to reach a more varied audience, including public and policymakers [16]. This emphasises the function of Twitter as a potential outreach tool, considerably increasing the overall reach of scientific messages [1].

Components and content of a tweet have a significant impact on the amplification of social media posts. In a recent study, the inclusion of a URL was identified as the most critical component in successful tweeting, whereas inclusion of an image was associated with less retweet success [6]. Although inclusion of a URL was also important in our study, this was only significant in the subgroup of users with over 1,000 followers. In comparison, the inclusion of multimedia and mention of other tweeter(s) were the only consistent independent predictors of a successful tweet in our study. Our component analysis in combination with subgroup analysis highlights that Twitter success and influence requires a strategic and tenacious performance, which is comparable with the findings from Cote et al. [16].

The most important limitation of this study relates to the intrinsic nature of social media analytics. For example, it was not possible to quantify the number of people who viewed the posts, nor to distinguish posts generated within or outside the congress hall or to identify the motivations for tweeting and retweeting, and not tweeting. We may have missed some of the

congress related posts which did not include the official congress hashtag. In addition, our aim was to determine the general characteristics of a successful tweet regardless of the topic; therefore, we have not themed the tweets by topic, and also the numbers in the subgroups were too small for adequately powered analysis. To expand this study and overcome some of the limitations, tweets and replies without the congress hashtag could be included manually and to allow this manual extraction the data could be extracted and shared immediately during the congresses.

In conclusion, social media could help disseminate scientific messages beyond the congress hall, if and when the posts constitute the right combination of components. Identifying ways to support more physicians and scientists to tweet original content should be explored.

Authors contributions

MC and GM conceptualised this study, GM performed the network analysis, DO performed the multivariable analysis. MC drafted the first and subsequent versions of the manuscript, and all authors provided critical feedback and contributed to the manuscript.

Financial support and sponsorship

None

Conflicts of interest

MC was a Steering Committee member of Trainee Association of ESCMID at the time of ECCMID 2018 and was actively involved in social media planning and implementation of

ECCMID2018. DO was a Steering Committee member of Trainee Association of ESCMID at the time of ECCMID 2017 and ECCMID 2018.

Acknowledgements

We would like to thank Miss Chantal Britt and Prof Winfried V. Kern who coordinated the social media activity during the ECCMID 2018, and active tweeters from the official ECCMID2018 Twitter handles; ESCMID and TAE accounts; Dr Selay Demirci, Dr Muge Cevik and Dr Zaira Palacios. We also would like to thank the ESCMID Executive Committee kindly providing permission to publish these findings. And lastly, we would like to thank all the active Twitter users during the Congress for making this activity a success.

References

1. Bonetta L. Should you be tweeting? Cell 2009; 139(3): 452-3.
2. Choo EK, Ranney ML, Chan TM, et al. Twitter as a tool for communication and knowledge exchange in academic medicine: A guide for skeptics and novices. Med Teach 2015; 37(5): 411-6.
3. Collins K, Shiffman D, Rock J. How Are Scientists Using Social Media in the Workplace? PLoS One 2016; 11(10): e0162680.
4. Soreide K, Mackenzie G, Polom K, Lorenzon L, Mohan H, Mayol J. Tweeting the meeting: Quantitative and qualitative twitter activity during the 38th ESSO conference. Eur J Surg Oncol 2018.
5. Bombaci SP, Farr CM, Gallo HT, et al. Using Twitter to communicate conservation science from a professional conference. Conserv Biol 2016; 30(1): 216-25.
6. Bruni LF, C; Giacomazzi, P. The Role of Multimedia Content in Determining the Virality of Social Media Information. Information 2012; 3(3): 2078-289.
7. Mitchell BG, Russo PL, Otter JA, Kiernan MA, Aveling L. What Makes a Tweet Fly? Analysis of Twitter Messaging at Four Infection Control Conferences. Infect Control Hosp Epidemiol 2017; 38(11): 1271-6.
8. ESCMID. Available at: https://www.escmid.org/membership_organization/about_escmid/. Accessed January 8.
9. SMRF. NodeXL. Available at: https://www.smrfoundation.org. Accessed January 6.
10. Mackenzie G. Monitoring health related activity on Twitter: Tips for individuals, organisations, campaigns and conference organisers. Available at:

https://scotpublichealth.com/2018/03/13/handing-over-the-reins-crowdsourcing-twitter-data-on-health-campaigns/. Accessed January 8.
11. Wikipedia. Centrality. Available at: https://en.wikipedia.org/wiki/Centrality Accessed January 6.
12. Mackenzie G. ECCMID 2017 Wakelet summary. Available at: https://wakelet.com/wake/129892a8-74a7-4087-9715-6d76b183f954. Accessed January 6.
13. Mackenzie G. ECCMID 2018 Wakelet Summary. Available at: https://wakelet.com/wake/52ecc4d0-b5fc-4ba2-9926-1cc55628dc63. Accessed January 6.
14. Smith MR, L; Shneiderman, B; Himelboim, I. Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters. Pew Research Center, 2014.
15. Hudson S, Mackenzie G. 'Not your daughter's Facebook': Twitter use at the European Society of Cardiology Conference 2018. Heart 2019; 105(2): 169.
16. Côté IM, Darling ES. Scientists on Twitter: Preaching to the choir or singing from the rooftops? FACETS 2018; 3(1): 682-94.

Table titles and legends

Table 1. Overview and comparison of social media activity at ECCMID 2017 and ECCMID 2018

Table 2. Factors associated with retweets in multivariable analysis

Supplementary table: Top 20 tweeters from #ECCMID2018 (by number of retweets)

Figure titles and legends:

Figure 1. Network activity of ECCMID 2018

Figure 1 Caption: This figure illustrates the extent of connections. Each picture represents a Twitter user who either tweeted, was mentioned, and/or retweeted posts using the #ECCMID2018 hashtag. The map displays a predominantly "tight crowd" pattern formed by highly connected tweeters (G1-G4).

Figure 2. Social media influencers tweeting, retweeting and mentioned in tweets using the #ECCMID2018 hashtag

Figure 2 Caption: This Venn diagram summarises all the users involved in ECCMID2018 social media activity (n=2,973 accounts). It classifies users into tweeters, retweeters and those mentioned and identifies most influential tweeters depicted in the center of the diagram.

Figure 3. Components of a successful tweet - univariable analysis

Figure 3 Caption: The components of a successful tweet were initially studied using univariable analysis based on ECCMID2018 data. Variables of interest were: inclusion of multimedia (images and/or video), a link to other information on the web (Uniform Resource Locator (URL)), mention of other tweeter(s), the number of followers of a Twitter user and use of hashtags other than the congress hashtag. There was a stepwise increase in the number of retweets received with the inclusion of more components.

Table 1: Overview and comparison of social media activity at ECCMID 2017 and ECCMID 2018

#	**ECCMID 2017**	**ECCMID 2018**	**p-value**
Total accounts tweeting, retweeting and/or mentioned in tweets	2,150	2,973	-
Number of active tweeters	590 (27%)	591 (20%)	<0.001
Number of accounts just tweeted	276 (13%)	214 (7%)	<0.001
Number of all accounts RTed	1,546 (72%)	2,389 (80%)	<0.001
Number of accounts just RTed	1,167 (54%)	1,904 (64%)	<0.001
Number of accounts mentioned	558 (26%)	655 (22%)	<0.001
Outcomes			
Number of tweets	4,657	4,213	-
Number of retweets	7,818	12,109	-

358 **Table 2. Factors associated with retweets in multivariable analysis**
359

Tweet characteristic	All tweeters	All tweeters (with inclusion of interaction terms)	Subgroup analysis <1,000 followers	Subgroup analysis >1,000 followers
	Odds ratio (95% C.I.)	Odds ratio (95% C.I.)	Odds ratio (95% C.I.)	Odds ratio (95% C.I.)
Inclusion of media (video or picture)	1.50 (1.30 – 1.74)	1.53 (1.33-1.78)	1.40 (1.14-1.72)	1.76 (1.43-2.17)
Mention of other tweeters	2.01 (1.75 – 2.32)	2.19 (1.84-2.60)	2.13 (1.76-2.57)	1.67 (1.35-2.07)
Inclusion of other hashtags	1.23 (1.07 – 1.42)	1.20 (1.04-1.39)	1.00 (0.82-1.22)	1.66 (1.35-2.04)
Inclusion of URL	1.42 (1.16 – 1.74)	2.32 (1.77-3.03)	1.28 (0.89-1.83)	1.74 (1.37-2.21)
The number of followers (per 100 increase)	1.01 (1.01 – 1.01)	1.02 (1.02-1.03)	N.A.	N.A.

360
361
362